

APR 23 2009

Serial No. 09/862,437

IN THE CLAIMS:

The text of all pending claims, (including withdrawn claims) is set forth below. Cancelled and not entered claims are indicated with claim number and status only. The claims as listed below show added text with underlining and deleted text with ~~striketrough~~.

1. (currently amended) A document reading apparatus presenting a plurality of documents designated as reading documents by a user, comprising:

a thematic hierarchy recognizing device recognizing a thematic hierarchy of each of the plurality of documents:

~~an inclusive relationship of topics from large topics to small topics by calculating by repeating a process for detecting a set of thematic boundaries in each layer of the thematic hierarchy, with each of a plurality of different window widths, wherein each of the thematic boundaries is detected based on a lexical cohesion score of description parts obtained from a similarity of vocabularies that appear in two adjacent windows with each of the window widths at each location in each of the plurality of documents, based on a vocabulary appearing in the adjacent windows; and~~

~~a thematic boundary based on the lexical cohesion score for a plurality of different window widths by repeating the calculating; and~~

~~a thematic hierarchy of each of the plurality of documents corresponding to the inclusive relationship of topics;~~

by correlating first and second thematic boundaries locating closely and detected with smaller and larger window widths, respectively;

a topic extracting device extracting a topic that commonly appears in the plurality of documents based on the thematic hierarchies recognized; and

a topic relation presenting device taking out a description part corresponding to the topic extracted from each of the plurality of documents and outputting the description parts taken out as related passages among said plurality of documents, wherein

regarding a topic set that comprises topics of various grading in the thematic hierarchies recognized, the topic extracting device calculates a relevance score between topics of the topic set based on lexical similarity of description parts corresponding to each topic of the topic set, and extracts a topic set having a relevance score equal to or more than a threshold that is set based on inclusive relationship of topics, and

the threshold, corresponding to any topic among topics which constitutes a target extraction of topic set, is a maximum value of calculated relevance score related to a topic which

Serial No. 09/862,437

is included in a subtree in thematic hierarchies.

2. (canceled)

3. (previously presented) The document reading apparatus according to claim 1, wherein the topic relation presenting device presents the description parts taken out side by side.

4. (previously presented) The document reading apparatus according to claim 3, wherein the topic relation presenting device presents the related parts and original documents in two windows, one of the windows including the related parts side by side and the other including the original documents side by side.

5. (previously presented) The document reading apparatus according to claim 3, wherein the topic relation presenting device presents summaries of the related parts.

6. (previously presented) The document reading apparatus according to claim 5, wherein the topic relation presenting device presents summaries of the related parts and original documents in two windows, one of the windows including the summaries side by side and the other including the original documents side by side.

7. (previously presented) The document reading apparatus according to claim 3, wherein the topic relation presenting device presents a plurality of thematic hierarchies corresponding to the plurality of documents and a correspondence relationship between the plurality of thematic hierarchies based on the plurality of common topics in a drawing, and presents a designated part of the plurality of documents in accordance with an instruction from the user given on the drawing.

8. (previously presented) The document reading apparatus according to claim 1, wherein the topic relation presenting device sets one document among the plurality of documents as a reference document, produces a new integrated document by merging contents of the reference document with description parts of another document related to the reference document, and outputs the integrated document.

Serial No. 09/862,437

9. (currently amended) A computer-readable storage medium encoded with a program for a computer that presents a plurality of documents designated as reading documents by a user, the program upon execution causing the computer to perform a method comprising:

recognizing a thematic hierarchy of each of the plurality of documents:

~~an inclusive relationship of topics from large topics to small topics by calculating by repeating a process for detecting a set of thematic boundaries in each layer of the thematic hierarchy, with each of a plurality of different window widths, wherein each of the thematic boundaries is detected based on a lexical cohesion score of description parts obtained from a similarity of vocabularies that appear in two adjacent windows with each of the window widths at each location in each of the plurality of documents, based on a vocabulary appearing in the adjacent windows; and a thematic boundary based on the lexical cohesion score for a plurality of different window widths by repeating the calculating; and a thematic hierarchy of each of the plurality of documents corresponding to the inclusive relationship of topics;~~

by correlating first and second thematic boundaries locating closely and detected with smaller and larger window widths, respectively;

extracting a topic that commonly appears in the plurality of documents based on the thematic hierarchies recognized; and

taking out a description part corresponding to the topic extracted from each of the plurality documents and outputting the description parts taken out as related passages among said plurality of documents, wherein

regarding a topic set that comprises topics of various grading in the thematic hierarchies recognized, a relevance score between topics of the topic set based on lexical similarity of description parts corresponding to each topic of the topic set is calculated, and a topic set having a relevance score equal to or more than a threshold that is set based on inclusive relationship of topics is extracted, and

the threshold, corresponding to any topic among topics which constitutes a target extraction of topic set, is a maximum value of calculated relevance score related to a topic which is included in a subtree in thematic hierarchies.

10. (canceled)

11. (currently amended) A document presenting method of presenting a plurality of documents designated as reading documents by a user, comprising:

recognizing a thematic hierarchy of each of the plurality of documents:

Serial No. 09/862,437

~~an inclusive relationship of topics from large topics to small topics by calculating by repeating a process for detecting a set of thematic boundaries in each layer of the thematic hierarchy, with each of a plurality of different window widths, wherein each of the thematic boundaries is detected based on a lexical cohesion score of description parts obtained from a similarity of vocabularies that appear in two adjacent windows with each of the window widths at each location in each of the plurality of documents, based on a vocabulary appearing in the adjacent windows; and~~

~~a thematic boundary based on the lexical cohesion score for a plurality of different window widths by repeating the calculating; and a thematic hierarchy of each of the plurality of documents corresponding to the inclusive relationship of topics;~~

~~by correlating first and second thematic boundaries locating closely and detected with smaller and larger window widths, respectively;~~

extracting a topic that commonly appears in the plurality of documents based on the thematic hierarchies recognized; and

taking out a description part corresponding to the topic extracted from each of the plurality documents and outputting the description parts taken out as related passages among said plurality of documents, wherein

regarding a topic set that comprises topics of various grading in the thematic hierarchies recognized, a relevance score between topics of the topic set based on lexical similarity of description parts corresponding to each topic of the topic set is calculated, and a topic set having a relevance score equal to or more than a threshold that is set based on inclusive relationship of topics is extracted, and

the threshold, corresponding to any topic among topics which constitutes a target extraction of topic set, is a maximum value of calculated relevance score related to a topic which is included in a subtree in thematic hierarchies.

12. (currently amended) A document reading apparatus presenting a plurality of documents designated as reading documents by a user, comprising:

thematic hierarchy recognizing means for recognizing a thematic hierarchy of each of the plurality of documents:

~~an inclusive relationship of topics from large topics to small topics by calculating by repeating a process for detecting a set of thematic boundaries in each layer of the thematic hierarchy, with each of a plurality of different window widths, wherein each of the thematic boundaries is detected based on a lexical cohesion score of description parts obtained from a~~

Serial No. 09/862,437

similarity of vocabularies that appear in two adjacent windows with each of the window widths at each location in each of the plurality of documents, based on a vocabulary appearing in the adjacent windows; and

a thematic boundary based on the lexical cohesion score for a plurality of different window widths by repeating the calculating; and

a thematic hierarchy of each of the plurality of documents corresponding to the inclusive relationship of topics;

by correlating first and second thematic boundaries locating closely and detected with smaller and larger window widths, respectively;

topic extracting means for extracting a topic that commonly appears in the plurality of documents based on the thematic hierarchies recognized; and

topic relation presenting means for taking out a description part corresponding to the topic extracted from each of the plurality documents and outputting the description parts taken out as related passages among said plurality of documents, wherein

regarding a topic set that comprises topics of various grading in the thematic hierarchies recognized, a relevance score between topics of the topic set based on lexical similarity of description parts corresponding to each topic of the topic set is calculated, and a topic set having a relevance score equal to or more than a threshold that is set based on inclusive relationship of topics is extracted, and

the threshold, corresponding to any topic among topics which constitutes a target extraction of topic set, is a maximum value of calculated relevance score related to a topic which is included in a subtree in thematic hierarchies.

13. (original) The document reading apparatus according to claim 1, wherein the thematic hierarchy recognizing device determines the thematic hierarchy according to topic-subtopic relations between topics.